IMMUNOGENIC HIV-1 MULTI-CLADE, MULTIVALENT CONSTRUCTS AND METHODS OF THEIR USE

PRIORITY CLAIM

This application claims the benefit of U.S. Provisional Patent Application No. 60/458,880 filed March 28, 2003, which is incorporated herein by reference in its entirety.

STATEMENT OF GOVERNMENT SUPPORT

This invention was made by the Centers for Disease Control and Prevention, an agency of the United States Government. Therefore, the U.S. Government has certain rights in this invention.

FIELD

This disclosure relates to compositions for induction of immune responses in vertebrates. More particularly, it relates to highly effective, broad spectrum multivalent constructs, both protein and nucleic acid, for inducing an immune response to an immunodeficiency virus, such as HIV-1. The disclosure further relates to vaccines comprising immunogenic compounds.

BACKGROUND

Vertebrates have developed a sophisticated system to protect themselves against a wide variety of hazards including various viruses and microorganisms, such as bacteria and fungi, as well as genetic diseases, neoplasia, and effects of a variety of toxins. The system has evolved based on the ability to recognize self as distinct from non-self or "foreign." A broad panoply of defense mechanisms are involved, including phagocytosis, lysis, such as complement mediated or perforin mediated lysis, and killer cells, such as cytotoxic T-lymphocytes (CTLs; also known as cytotoxic/suppressor T-cells, Tc/s), natural killer cells, antibody dependent cytotoxic cells, and the like. Various cell types offer different mechanisms whereby the invader or endogenous diseased cell may be eliminated.

A key to the immune defensive mechanism is the T-cell. For instance, it is well known that the adaptive immune system shows a much stronger response on second, as compared to first, encounter with an antigen. This phenomenon is exploited in vaccination, which works by inducing a state of lasting immunity known as immunological memory. Immunological memory requires activation of T-lymphocytes specific for the vaccine antigen.

T-cells have been found to be "restricted" in that they respond to an antigen in relation to one or a few specific molecules (now called major histocompatibility or MHC molecules) associated with their natural host. *In vitro*, T-cells from a host of one haplotype respond to an antigen in relation to an MHC molecule of a different haplotype host. The T-cell receptor recognition repertoire appears to be narrower than the recognition repertoire of immunoglobulins produced by B-cells. In addition, rather than directly binding to an antigen as do antibodies and other immunoglobulins, the T-cell receptor appears to require concomitant binding to a foreign antigen and an MHC molecule.

MHC molecules are divided into two classes, Class I and Class II. The former class is relatively ubiquitous on vertebrate cells, while the latter is generally limited to lymphocytes, macrophages, and dendritic cells. Functionally different T-cells appear to be activated in relation to one or the other class of MHC molecules. The nature of the activity of a T-cell varies with the Class of the MHC molecule to which it is complementary. A T-cell clone recognizes a specific antigen in conjunction with a specific MHC allele. Furthermore, variation in the antigen structure affects the nature of the response when the T-cell, antigen, and antigen presenting cell are brought together. Depending upon the nature of the structural change, three possibilities are encountered: no change, increased stimulation or decreased stimulation of an immune response to the antigen.

T-lymphocytes detect foreign polypeptide antigens by recognizing – via the T-cell receptor ("TCR") – peptide fragments derived from the antigen. Most T-lymphocytes, however, are MHC restricted, that is, they recognize only complexes of peptides bound to the highly polymorphic membrane proteins encoded by Class I and Class II MHC genes and presented (displayed) on the surface of an accessory cell (designated an antigen-presenting cell or "APC"), in which the antigen has been processed.

Antigens can be processed by one of two pathways, depending on their origin, inside or outside the APC. In a first pathway, foreign material from outside a cell is engulfed by a specialized APC (often a macrophage or B-cell), which breaks down the material and complexes the processed antigen with Class II MHC molecules. In particular, MHC Class II molecules are synthesized in the endoplasmic reticulum with their antigenic peptide binding sites blocked by the invariant chain protein (Ii). These MHC Class II-Ii protein complexes are transported from the endoplasmic reticulum to a post-Golgi compartment where Ii is released by proteolysis and a specific antigenic peptide becomes bound to the MHC Class II molecule.

Class II MHC molecules are expressed primarily on cells involved in initiating and sustaining immune responses, such as T-lymphocytes, B-lymphocytes, and macrophages. Complexes of Class II MHC molecules and immunogenic peptides are recognized by helper T-lymphocytes (also known as helper/accessory T-cells, "Th") and induce proliferation of Th lymphocytes. Class II MHC complexes also stimulate secretion of cytokines by Th cells, resulting in amplification of the immune response to the particular immunogenic peptide that is displayed. Th1 cells produce interferon-γ and other cytokines that stimulate CTLs, while other cytokines produced by Th2 cells help B-cells to produce antibodies.

A second antigen processing pathway is generally involved with foreign or aberrant proteins made within cells, such as virus-infected or malignant cells. These proteins are subjected to partial proteolysis by the proteosome within such cells, so as to form peptide fragments that then associate with Class I MHC molecules and are transported to the cell surface for presentation to T-cells. Class I MHC molecules are expressed on almost all nucleated cells, and complexes of Class I MHC molecules and bound immunogenic peptides are recognized by CTLs, which then destroy the antigen-bearing cells. CTLs are particularly important in tumor rejection and in fighting viral infections.

For a CTL to recognize an antigen in the form of a peptide fragment bound to the MHC class I molecule, that antigen must normally be endogenously synthesized by the cell and a portion degraded into small peptide fragments in the cytoplasm. Some of these small peptides translocate into a pre-Golgi compartment and interact with Class I heavy chains to facilitate proper folding and association with the subunit ß2 microglobulin. The peptide-MHC Class I complex is then routed to the cell surface for expression and potential recognition by specific CTLs.

By these dual antigen processing pathways, appropriate defenses are generated against both exogenous and internally produced antigens. Thus, antigens taken up from the extracellular environment eventually elicit B-cells to produce antibodies that protect the organism against a subsequent challenge by an agent comprising the exogenous antigen. On the other hand, antigens comprised of abnormal structures made within an abnormal or errant cell (for example a virus-infected or malignant cell) activate an immune response that eventually leads to killing of the errant cell. There is considerable interest in methods for better stimulating immune responses to antigens that are processed by either of these two pathways and presented by either MHC Class I or Class II molecules.

In view of the above knowledge, it is understandable that there has been substantial interest in using short peptides to affect an immune response *in vivo* and *in vitro*, to provide stimulation or inactivation of a particular response. Thus, appropriate immunogenic peptides might modulate a natural immune response to a particular event, either by activating particular lymphocytes to enhance a protective response or by deactivating particular lymphocytes to diminish or prevent an undesirable response.

The human immunodeficiency virus (HIV-1, also referred to as HTLV-III, LAV or HTLV-III/LAV) is the etiological agent of the acquired immune deficiency syndrome (AIDS) and related disorders (see, for example, Barre-Sinoussi *et al.*, *Science* 220:868-871, 1983; Gallo *et al.*, *Science* 224:500-503, 1984; Levy *et al.*, *Science* 225:840-842, 1984; Siegal *et al.*, *N. Engl. J. Med.* 305:1439-1444, 1981). AIDS patients usually have a long asymptomatic period followed by the progressive degeneration of the immune system and the central nervous system. Replication of the virus is highly regulated, and both latent and lytic infection of the CD4 positive helper subset of T-lymphocytes occur in tissue culture (Zagury *et al.*, *Science* 231:850-853, 1986). Molecular studies of HIV-1 show that it encodes a number of genes (Ratner *et al.*, *Nature* 313:277-284, 1985; Sanchez-Pescador *et al.*, *Science* 227:484-492, 1985), including three structural genes – gag, pol and env – that are common to all retroviruses. Nucleotide sequences from viral genomes of other retroviruses, particularly HIV-2 and simian immunodeficiency viruses (SIV; previously referred to as STLV-III), also contain these structural genes (Guyader *et al.*, *Nature* 326:662-669, 1987; Chakrabarti *et al.*, *Nature* 328:543-547, 1987).

Development of an effective HIV vaccine is a major challenge due to antigenic variation and immune escape mechanisms. Strategies that include the use of recombinant DNA technology and novel antigen delivery methods are being applied to the development of HIV vaccines. Most HIV-1

vaccine constructs (DNA and recombinant protein vaccine) are subtype-specific and designed to prime only one arm of the immune system, that is, CTL responses or humoral B-cell responses. Emerging data suggest that broadly reactive T-cell responses, as well as neutralizing antibody responses are likely to be required for an effective immune response against HIV-1. Additionally, current human phase III vaccine trials using recombinant envelope proteins, suggest that immunity to HIV-1 envelope proteins is probably not sufficient for complete protection against HIV-1. Thus the results from multiple studies suggest that additional epitopes as well as activation of both arms of the immune system may be required for an effective HIV-1 vaccine.

By way of one example of peptide immunogens, Peter *et al.* (*Vaccine* 19:4121-4129, 2001) disclose induction of a CTL response against multiple CTL epitopes present in HIV proteins using short synthetic peptides. Four HLA-A2.1 restricted peptides (RT 476-484, p17 77-85, gp41 814-823, RT 956-964) that showed stable binding to the HLA-A2.1 molecule in an *in vitro* binding assay were able to elicit a strong specific immune response in HLA-A2.1 transgenic mice when injected with a peptide ("P30") used as a universal T-cell helper epitope, in incomplete Freund adjuvant (IFA) or a nonionic emulsifier (Montanide™ ISA 720). The use of biodegradable poly-L-glutamic acid (PLGA) microspheres (MS) as adjuvant was also successfully tested for all peptides.

Many studies of cross-clade recognition of HIV epitopes have been carried out (see, for example, Wilson *et al.*, *AIDS Res. Hum. Retroviruses* 14:925-937, 1998; McAdam *et al.*, *AIDS* 12:571-579, 1998; Lynch *et al.*, *J Infect Dis.* 178:1040-1046, 1998; Boyer *et al. Dev. Biol. Stand.* 95:147-53, 1998; Cao *et al.*, *J. Virol.* 71:8615-8623, 1997; Durali *et al.*, *Virol.* 72:3547 3553, 1998). These studies often used whole-gene, vaccinia-expressed constructs to probe CTL lines from HIV-1 infected or HIV-1 vaccinated volunteers for CTL responses. What appeared to be cross-clade recognition by CTLs in these experiments may have been recognition of CTL epitopes that are conserved within the large gene constructs cloned into the vaccinia constructs and into the vaccine strain (or the autologous strain). Where responses to specific peptides, and their altered sequences in other HIV strains, have been tested, and the peptides have been mapped, some studies have shown a lack of cross-strain recognition (Dorrel *et al.*, HIV Vaccine Development Opportunities And Challenges Meeting, Abstract 109 (Keystone, Colo., January 1999)). Studies of virus escape from CTL recognition carried out on HIV-1 infected individuals have also shown that viral variation at the amino acid level may abrogate effective CTL responses (Koup, *J. Exp. Med.* 180:779-782, 1994; Dai *et al.*, *J. Virol.* 66:3151-3154, 1992; Johnson *et al.*, *J. Exp. Med.* 175:961-971, 1992).

Unfortunately, existing candidate HIV-1 vaccines are subtype specific, and are expected not protect against diverse natural HIV-1 infections. This is true of both DNA vaccine constructs as well as recombinant protein vaccines. Furthermore, many of the existing constructs have focused on priming only one arm of the immune system, that is, cell mediated T-cell responses or humoral B-cell responses. In addition, while some DNA constructs have shown promising results in lowering viremia in animal model systems, none has been able to confer sterilizing immunity. These data suggest that both B-cell and T-cell responses may be needed for a protective immune response

against HIV-1. Additionally, current human phase 3 vaccine trials using recombinant envelope proteins, suggests that immunity to HIV-1 envelope proteins is probably not sufficient for complete protection against HIV-1. Prime-boost strategy using recombinant envelope from HIV-1 subtype B also has not been successful in boosting the immune responses.

As the HIV epidemic continues to spread world wide, the need for effective immune-stimulatory compositions and vaccines remains urgent.

SUMMARY OF THE DISCLOSURE

Multi-clade multivalent (MCMV) (polyepitope; multi-epitope) polypeptides and mixtures of polypeptides have been developed, which can be used to stimulate immune responses to HIV-1 in vertebrates. In various embodiments, these polypeptides and polypeptide mixtures include immunogenic CTL, T- and/or B-cell determinants that are capable of eliciting broad and effective immune responses against diverse subtypes of HIV-1. Immunogens described herein are designed to be subtype-independent and will provide both prime and boost reagents for worldwide use.

Also described herein are recombinant MCMV constructs that can be used directly or indirectly to protect subjects against infection by multiple HIV-1 subtypes. These constructs are designed to elicit T-cell, B-cell, or both T-cell and B-cell responses against highly conserved epitopes within multiple HIV-1 subtypes. The constructs, when integrated into a vector, can be used as immunogens, can be used as DNA vaccines, and can be used as sources of recombinant protein for stimulation of immune responses in subjects, as well as for protein boosts to subjects who have received a nucleic acid construct previously.

Without being bound by theory, it is believed that the MCMV HIV-1 constructs and polypeptides provide universal immune stimulants and vaccines, capable of effective use in any part of the world affected by the HIV-1 epidemic.

The construction and design of specific provided constructs are particularly useful in that they allow convenient addition/deletion of epitopes, and contain specific cellular targeting domains that optimize antigen processing and recognition.

The provided constructs and proteins encoded thereby also can be combined with other epitope-based constructs to generate, for instance, multi-pathogen vaccines.

The foregoing and other features and advantages will become more apparent from the following detailed description of several embodiments, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a series of schematic drawings of embodiments of specific multi-clade, multivalent gene constructs.

FIG 1A shows one embodiment of a CTL-stimulating MCMV (MCMV-CTL) construct. Epitopes were chosen based on prior responses of HIV+ individuals, predicted HLA binding and

sequence conservation among multiple HIV subtypes. In the Examples provided below, such a gene construct was assembled using synthetic single stranded oligonucleotides (100-130 mers) that contain strings of 3-6 CTL epitopes and linker amino acid sequences (exemplified by the tri-amino acid KAA), which were included to improve processing of epitopes. A modified human ubiquitin peptide is optionally added to the amino terminus of the molecule to further increase CTL epitope processing.

FIG 1B shows one embodiment of a MCMV-AB/Th construct. Antibody and T-helper epitopes conserved among multiple subtypes of HIV-1 were chosen and single stranded oligos (100-120 mers) for these epitopes were synthesized. The lysosomal integral membrane protein-II (LIMP-II) signal sequence is optionally included to enhance processing of T-helper epitopes.

FIG 1C is a schematic illustration of a MCMV-CTL, illustrating that the same nucleic acid construct can be used to generate both ubiquitin+ and ubiquitin- sequences by differential placement of the forward primer used to amplify the sequence.

FIG 1D shows an alternative embodiment, in which both the CTL and the AB/Th epitopes are provided in the same recombinant construct. In a combined construct such as this, the order of the different epitope sets can be rearranged.

Figure 2 is a schematic illustration of the assembly of representative MCMV construct. In the illustrated embodiment (FIG 2A), overlapping single stranded oligonucleotides (100-130 mers), spanning the full length of the MCMV-CTL-ubiquitin construct (1.5 kb) were synthesized (eight forward and eight reverse). Through a series of splicing overlap extension (SOE), polymerase chain reaction (PCR) and cloning steps, a 1,553 base pair recombinant nucleic acid sequence was generated and then cloned into pVax-1 (Invitrogen, Carlsbad, CA) (FIG 2B). Alternatively, the construct was also assembled without including ubiquitin (FIG 2C).

Figure 3 is an schematic drawing of vector PTriex-4 (Novagen, Madison, WI), which contains a representative MCMV-CTL construct, which can be used for production of recombinant protein in either bacterial, mammalian or insect cells.

Figure 4 is a western blot analysis showing expression of a MCMV-CTL-ubiquitin polypeptide fusion construct in *E. coli*. The fusion protein is predicted to be 64 kDa (57 kDa plus the 7 kDa expression tag); expressed protein is indicated in the right hand most lane.

Figure 5 is a western blot analysis of extract from HeLa cells transfected with pVax-1 (Invitrogen, Carlsbad, CA) containing the MCMV-CTL-ubiquitin sequence, using an anti-ubiquitin antibody for detection. Cells were transfected using various concentrations of GeneJuice reagent (Merck Biosciences, San Diego, CA) with 1 μg of DNA. Cells were harvested 24 and 48 hours post transfection. A ubiquitinated protein of the correct predicted molecular weight (~56 kDa) of the synthetic MCMV-CTL-ubiquitin construct is clearly visible, as is normal cellular ubiquitin (~10 kDa).

Figure 6 is a series of bar graphs demonstrating breadth and magnitude of CTL responses observed to peptides contained in the MCMV-CTL construct with peripheral blood mononuclear cells

(PBMCs) from individuals chronically infected with HIV-1 subtype B. The responses are reported as spot forming cells per 10^5 PBMCs.

Figure 7 is a bar graph demonstrating percentage predicted epitope recognition based on subject HLA type. The percentage of predicted epitopes that were targeted by patients' $CD8^+$ cells in the Elispot assay is shown.

SEQUENCE LISTING

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids, as defined in 37 C.F.R. 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand. In the accompanying sequence listing:

SEQ ID NO: 1 shows the nucleic acid sequence and amino acid sequence of MCMV-CTL-ubiquitin. The nucleic acid sequence includes unique restriction sites at positions 6-11 and 1548-1553. These restriction sites can be used to insert the epitope construct into different vectors.

SEQ ID NO: 2 shows the amino acid sequence of MCMV-CTL-ubiquitin. Ubiquitin is positions 1-76. The "KAA" spacer peptide appears (amino acid positions 117-119, 169-172, 242-244, 288-290, 317-319, 367-369, 417-419, and 466-468) throughout the remainder of the sequence between strings of three to five CTL epitopes.

SEQ ID NO: 3 shows the nucleic acid sequence and amino acid sequence of MCMV-CTL (no ubiquitin). The nucleic acid sequence includes unique restriction sites at positions 1-6 and 1318-1323. These restriction sites can be used to insert the epitope construct into different vectors.

SEQ ID NO: 4 shows the amino acid sequence of MCMV-CTL (no ubiquitin). The "KAA" spacer appears throughout the sequence, at positions analogous to those in SEQ ID NO: 2.

SEQ ID NO: 5 shows the amino acid sequence of CTLUbiquitinNC (without mouse and monkey control epitopes).

SEQ ID NO: 6 shows the amino acid sequence of CTLNC (no ubiquitin, without mouse and monkey control epitopes).

SEQ ID NO: 7 shows the nucleic acid sequence and amino acid sequence of MCMV-AB/Th with LIMP-II.

SEQ ID NO: 8 shows the amino acid sequence of MCMV-AB/Th with LIMP-II.

SEQ ID NO: 9 shows the nucleic acid sequence and amino acid sequence of MCMV-AB/Th without LIMP-II.

SEQ ID NO: 10 shows the amino acid sequence of MCMV-AB/Th without LIMP-II.

SEQ ID NOs: 11-22 show the amino acid sequences of additional HIV-1 CTL antigenic fragments/epitopes.

SEQ ID NOs: 23-45 show the amino acid sequences of control peptides.

SEQ ID NOs: 46-59 show the amino acid sequences of additional HIV-1 B-cell antigenic fragments/epitopes.

SEQ ID NOs: 60-64 show the amino acid sequences of additional HIV-1 T-helper cell antigenic fragments/epitopes.

DETAILED DESCRIPTION

I. Abbreviations

HIV	human immunodeficiency virus
LIMP-II	lysosomal integral membrane protein II
MCMV	multi-clade multivalent
MCMV-AB/Th	B-cell/T-cell epitopes MCMV construct/polypeptide
MCMV-CTL	CTL epitopes MCMV construct/polypeptide
PCR	polymerase chain reaction
SOE	splicing overlap extension

II. Terms

Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes VII*, published by Oxford University Press, 2000 (ISBN 019879276X); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Publishers, 1994 (ISBN 0632021829); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by Wiley, John & Sons, Inc., 1995 (ISBN 0471186341); and other similar references.

In order to facilitate review of the various embodiments of the invention, the following explanations of specific terms are provided:

Adjuvant: A substance that non-specifically enhances the immune response to an antigen. Development of vaccine adjuvants for use in humans is reviewed in Singh *et al.*, *Nat. Biotechnol.* 17:1075-1081, 1999, which discloses that, at the time of its publication, aluminum salts and the MF59 microemulsion are the only vaccine adjuvants approved for human use.

Binding or stable binding (of an oligonucleotide): An oligonucleotide binds or stably binds to a target nucleic acid if a sufficient amount of the oligonucleotide forms base pairs or is hybridized to its target nucleic acid, to permit detection of that binding. Binding can be detected by either physical or functional properties of the target:oligonucleotide complex. Binding between a target and an oligonucleotide can be detected by any procedure known to one skilled in the art, including both functional and physical binding assays. Binding may be detected functionally by determining whether binding has an observable effect upon a biosynthetic process such as expression of a gene, DNA replication, transcription, translation and the like.

Physical methods of detecting the binding of complementary strands of DNA or RNA are well known in the art, and include such methods as DNase I or chemical footprinting, gel shift and

affinity cleavage assays, Northern blotting, dot blotting and light absorption detection procedures. For example, one method that is widely used, because it is so simple and reliable, involves observing a change in light absorption of a solution containing an oligonucleotide (or an analog) and a target nucleic acid at 220 to 300 nm as the temperature is slowly increased. If the oligonucleotide or analog has bound to its target, there is a sudden increase in absorption at a characteristic temperature as the oligonucleotide (or analog) and target disassociate from each other, or melt.

The binding between an oligomer and its target nucleic acid is frequently characterized by the temperature (T_m) at which 50% of the oligomer is melted from its target. A higher (T_m) means a stronger or more stable complex relative to a complex with a lower (T_m).

Complementarity and percentage complementarity: Molecules with complementary nucleic acids form a stable duplex or triplex when the strands bind, (hybridize), to each other by forming Watson-Crick, Hoogsteen or reverse Hoogsteen base pairs. Stable binding occurs when an oligonucleotide remains detectably bound to a target nucleic acid sequence under the required conditions.

Complementarity is the degree to which bases in one nucleic acid strand base pair with the bases in a second nucleic acid strand. Complementarity is conveniently described by percentage, that is, the proportion of nucleotides that form base pairs between two strands or within a specific region or domain of two strands. For example, if 10 nucleotides of a 15-nucleotide oligonucleotide form base pairs with a targeted region of a DNA molecule, that oligonucleotide is said to have 66.67% complementarity to the region of DNA targeted.

A thorough treatment of the qualitative and quantitative considerations involved in establishing binding conditions that allow one skilled in the art to design appropriate oligonucleotides for use under the desired conditions is provided by Beltz *et al.*, *Methods Enzymol.* 100:266-285, 1983, and by Sambrook *et al.* (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

DNA (deoxyribonucleic acid): DNA is a long chain polymer which comprises the genetic material of most living organisms (some viruses have genes comprising ribonucleic acid (RNA)). The repeating units in DNA polymers are four different nucleotides, each of which comprises one of the four bases, adenine, guanine, cytosine and thymine bound to a deoxyribose sugar to which a phosphate group is attached. Triplets of nucleotides (referred to as codons) code for each amino acid in a polypeptide, or for a stop signal. The term codon is also used for the corresponding (and complementary) sequences of three nucleotides in the mRNA into which the DNA sequence is transcribed.

Unless otherwise specified, any reference to a DNA molecule is intended to include the reverse complement of that DNA molecule. Except where single-strandedness is required by the text herein, DNA molecules, though written to depict only a single strand, encompass both strands of a double-stranded DNA molecule. Thus, a reference to the nucleic acid molecule that encodes a particular MCMV construct, or a fragment thereof, encompasses both the sense strand and its reverse

complement. Thus, for instance, it is appropriate to generate probes or primers from the reverse complement sequence of the disclosed nucleic acid molecules.

Deletion: The removal of a sequence of DNA, the regions on either side of the removed sequence being joined together. Similar, this term can refer to the removal (for example, though genetic engineering means) of an amino acid sequence within a protein, the regions on either side of the removed sequence being joined together.

Epitope tags : Short stretches of amino acids to which a specific antibody can be raised, which in some embodiments allows one to specifically identify and track the tagged protein that has been added (for instance) to a living organism or to cultured cells. Detection of the tagged molecule can be achieved using a number of well known techniques. Examples of such techniques include (but are not limited to): immunohistochemistry, immunoprecipitation, flow cytometry, immunofluorescence microscopy, ELISA, immunoblotting ("western" blotting), and affinity chromatography. Examples of well known epitope tags include FLAG, T7, HA (hemagglutinin) and myc. The FLAG tag (DYKDDDDK) is beneficially used in some embodiments because high quality reagents are available to be used for its detection.

Hybridization: Oligonucleotides and their analogs hybridize by hydrogen bonding, which includes Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary bases. Generally, nucleic acid consists of nitrogenous bases that are either pyrimidines (cytosine (C), uracil (U), and thymine (T)) or purines (adenine (A) and guanine (G)). These nitrogenous bases form hydrogen bonds between a pyrimidine and a purine, and the bonding of the pyrimidine to the purine is referred to as "base pairing." More specifically, A will hydrogen bond to T or U, and G will bond to C. "Complementary" refers to the base pairing that occurs between to distinct nucleic acid sequences or two distinct regions of the same nucleic acid sequence.

"Specifically hybridizable" and "specifically complementary" are terms that indicate a sufficient degree of complementarity such that stable and specific binding occurs between the oligonucleotide (or its analog) and the DNA or RNA target. The oligonucleotide or oligonucleotide analog need not be 100% complementary to its target sequence to be specifically hybridizable. An oligonucleotide or analog is specifically hybridizable when binding of the oligonucleotide or analog to the target DNA or RNA molecule interferes with the normal function of the target DNA or RNA, and there is a sufficient degree of complementarity to avoid non-specific binding of the oligonucleotide or analog to non-target sequences under conditions where specific binding is desired, for example under physiological conditions in the case of *in vivo* assays or systems. Such binding is referred to as specific hybridization.

Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing nucleic acid sequences. Generally, the temperature of hybridization and the ionic strength (especially the Na^+ concentration) of the hybridization buffer will determine the stringency of hybridization, though wash times also influence stringency. Calculations regarding hybridization

conditions required for attaining particular degrees of stringency are discussed by Sambrook *et al.* (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, chapters 9 and 11, herein incorporated by reference.

For purposes of the present disclosure, "stringent conditions" encompass conditions under which hybridization will only occur if there is less than 25% mismatch between the hybridization molecule and the target sequence. "Stringent conditions" may be broken down into particular levels of stringency for more precise definition. Thus, as used herein, "moderate stringency" conditions are those under which molecules with more than 25% sequence mismatch will not hybridize; conditions of "medium stringency" are those under which molecules with more than 15% mismatch will not hybridize, and conditions of "high stringency" are those under which sequences with more than 10% mismatch will not hybridize. Conditions of "very high stringency" are those under which sequences with more than 6% mismatch will not hybridize.

Isolated: An "isolated" biological component (such as a nucleic acid molecule, protein or organelle) has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, that is, other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. Nucleic acids and proteins that have been "isolated" include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

Lysosomal compartment: Membrane-bound acidic vacuoles containing lysosomal-associated membrane protein (LAMP) molecules in the membrane, hydrolytic enzymes that function in antigen processing, and MHC class II molecules for antigen recognition and presentation. This compartment functions as a site for degradation of foreign materials internalized from the cell surface by any of a variety of mechanisms including endocytosis, phagocytosis and pinocytosis, and of intracellular material delivered to this compartment by specialized autolytic phenomena (de Duve, *Eur. J. Biochem.* 137:391, 1983).

The biosynthesis and vacuolar targeting mechanisms of the hydrolytic enzymes present in the lysosomal compartment have been extensively studied (Kornfeld & Mellman, *Ann. Rev. Cell Biol.*, 5:483, 1989). Newly synthesized hydrolases in the Golgi apparatus acquire mannose 6-phosphate groups that serve as specific recognition markers for the binding of these enzymes to mannose 6-phosphate receptors which are then targeted in some unknown manner to a prelysosomal vacuole. There the receptor-enzyme complex is dissociated by low pH, and the receptors recycle to the Golgi apparatus, while the enzyme-containing vacuole matures into a lysosome.

Studies of the structure and function of the lysosomal membrane were initiated in 1981 by August and colleagues with the discovery of major cellular glycoproteins that were subsequently termed LAMP-1 and LAMP-2 due to their predominant localization in the lysosomal membrane. Analogous proteins were subsequently identified in rat, chicken and human cells. Typically, LAMP-1, as deduced from a cDNA clone (Chen *et al.*, *J. Biol. Chem.*, 263:8754, 1988) consists of a

polypeptide core of about 382 amino acids (~42,000 Da) with a large (346-residue) intraluminal amino-terminal domain followed by a 24-residue hydrophobic transmembrane region and short (12-residue) carboxyl-terminal cytoplasmic tail. The intraluminal domain is highly glycosylated, being substituted with about 20 asparagine linked complex-type oligosaccharides and consists of two ~160-residue homology units that are separated by a proline/serine-rich region. Each of these homologous domains contains four uniformly spaced cysteine residues, disulfide bonded to form four 36-38-residue loops symmetrically placed within the two halves of the intraluminal domain (Arterburn *et al., J. Biol. Chem.*, 265:7419, 1990). The LAMP-2 molecule is highly similar to LAMP-1 in overall amino acid sequence (Cha *et al., J. Biol. Chem.*, 265:5008, 1990).

Another glycoprotein, described as CD63, MEA491 or LIMP-I, is also found in lysosomal membranes, as well as other in vacuolar structures (Azorza *et al., Blood*, 78:280, 1991). This molecule is distinctly different from the LAMPs, with a core polypeptide of about 25,000 Da and four transmembrane domains, but it has a cytoplasmic structure and sequence similar to the LAMP molecules. There is also extensive amino acid sequence similarity between this protein and a family of other molecules that also contain four membrane spanning domains, including the *Schistosoma mansoni* membrane protein SM23, CD37, the tumor-associated antigen CO-029, and the target of antiproliferative antibody-1.

LIMP-II is an additional glycoprotein present in the membrane of lysosomes and secretory granules with lysosomal properties (Vega *et al., J. Biol. Chem.*, 266:16818, 1991). A sequence near the amino-terminus with properties of an uncleavable signal peptide and a hydrophobic amine acid segment near the carboxyl end suggest that the protein is anchored in cell membranes at two sites by two short cytoplasmic tails at the amine and carboxyl-terminal ends of the protein. The molecule does not have sequence homology to any of the other described lysosomal membrane protein, but is highly similar to the cell surface protein CD36, which is involved in cell adhesion.

Nucleotide: This term includes, but is not limited to, a monomer that includes a base linked to a sugar, such as a pyrimidine, purine or synthetic analogs thereof, or a base linked to an amino acid, as in a peptide nucleic acid (PNA). A nucleotide is one monomer in a polynucleotide. A nucleotide sequence refers to the sequence of bases in a polynucleotide.

Oligonucleotide: An oligonucleotide is a plurality of joined nucleotides joined by native phosphodiester bonds, between about 6 and about 300 nucleotides in length. An oligonucleotide analog refers to moieties that function similarly to oligonucleotides but have non-naturally occurring portions. For example, oligonucleotide analogs can contain non-naturally occurring portions, such as altered sugar moieties or inter-sugar linkages, such as a phosphorothioate oligodeoxynucleotide. Functional analogs of naturally occurring polynucleotides can bind to RNA or DNA, and include PNA molecules.

Particular oligonucleotides and oligonucleotide analogs can include linear sequences up to about 200 nucleotides in length, for example a sequence (such as DNA or RNA) that is at least 6

bases, for example at least 8, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100 or even 200 bases long, or from about 6 to about 50 bases, for example about 10-25 bases, such as 12, 15 or 20 bases.

Operably linked: A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame.

Open reading frame: A series of nucleotide triplets (codons) coding for amino acids without any internal termination codons. These sequences are usually translatable into a peptide/polypeptide/protein.

Parenteral: Administered outside of the intestine, for example, not via the alimentary tract. Generally, parenteral formulations are those that will be administered through any possible mode except ingestion. This term especially refers to injections, whether administered intravenously, intrathecally, intramuscularly, intraperitoneally, or subcutaneously, and various surface applications including intranasal, intradermal, and topical application, for instance.

Pharmaceutically acceptable carriers: The pharmaceutically acceptable carriers useful with compositions described herein are conventional. Martin, *Remington's Pharmaceutical Sciences*, published by Mack Publishing Co., Easton, PA, 19th Edition, 1995, describes compositions and formulations suitable for pharmaceutical delivery of the nucleotides and proteins herein disclosed.

In general, the nature of the carrier will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (for example, powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically-neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

Probes and primers: A probe comprises an isolated nucleic acid attached to a detectable label or other reporter molecule. Typical labels include radioactive isotopes, enzyme substrates, co-factors, ligands, chemiluminescent or fluorescent agents, haptens, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, for example, in Sambrook *et al.* (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 and Ausubel *et al. Short Protocols in Molecular Biology*, 4th ed., John Wiley & Sons, Inc., 1999.

Primers are short nucleic acid molecules, for instance DNA oligonucleotides 10 nucleotides or more in length, for example that hybridize to contiguous complementary nucleotides or a sequence

to be amplified. Longer DNA oligonucleotides may be about 15, 20, 25, 30 or 50 nucleotides or more in length. Primers can be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then the primer extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, for example, by the PCR or other nucleic-acid amplification methods known in the art. Other examples of amplification include strand displacement amplification, as disclosed in U.S. Patent No. 5,744,311; transcription-free isothermal amplification, as disclosed in U.S. Patent No. 6,033,881; repair chain reaction amplification, as disclosed in WO 90/01069; ligase chain reaction amplification, as disclosed in EP-A-320 308; gap filling ligase chain reaction amplification, as disclosed in 5,427,930; and NASBA™ RNA transcription-free amplification, as disclosed in U.S. Patent No. 6,025,134.

Methods for preparing and using nucleic acid probes and primers are described, for example, in Sambrook *et al.* (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; Ausubel *et al. Short Protocols in Molecular Biology*, 4th ed., John Wiley & Sons, Inc., 1999; and Innis *et al. PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA, 1990. Amplification primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, MA). One of ordinary skill in the art will appreciate that the specificity of a particular probe or primer increases with its length. Thus, in order to obtain greater specificity, probes and primers can be selected that comprise at least 20, 25, 30, 35, 40, 45, 50 or more consecutive nucleotides of a target nucleotide sequences.

Protein: A biological molecule, particularly a polypeptide, expressed by a gene and comprised of amino acids.

Purified: The term "purified" does not require absolute purity (for example, the absence of all other substances); rather, it is intended as a relative term. Thus, for example, a purified protein preparation is one in which the protein referred to is more pure than the protein in its natural environment within a cell or within a production reaction chamber (as appropriate).

Recombinant: A recombinant nucleic acid is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination can be accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, for example, by genetic engineering techniques.

Specific binding agent: An agent that binds substantially only to a defined target. Thus a protein-specific binding agent binds substantially only the defined protein, or to a specific region within the protein. As used herein, a protein-specific binding agent includes antibodies and other agents that bind substantially to a specified polypeptide. The antibodies may be monoclonal or

polyclonal antibodies that are specific for the polypeptide, as well as immunologically effective portions ("fragments") thereof.

Antibodies may be produced using standard procedures described in a number of texts, including Harlow and Lane, *Using Antibodies: A Laboratory Manual*, CSHL, New York, 1999. The determination that a particular agent binds substantially only to the target protein may readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Harlow and Lane, *Using Antibodies: A Laboratory Manual*, CSHL, New York, 1999). Western blotting may be used to determine that a given target protein binding agent, such as a monoclonal antibody, binds substantially only to the specified target protein.

Shorter fragments of antibodies can also serve as specific binding agents. For instance, Fabs, Fvs, and single-chain Fvs (SCFvs) that bind to target protein (or epitope within a protein or fusion protein) would also be specific binding agents for that protein (or epitope). These antibody fragments are defined as follows: (1) FAb, the fragment which contains a monovalent antigen-binding fragment of an antibody molecule produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain; (2) FAb', the fragment of an antibody molecule obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two FAb' fragments are obtained per antibody molecule; (3) (FAb')2, the fragment of the antibody obtained by treating whole antibody with the enzyme pepsin without subsequent reduction; (4) F(Ab')2, a dimer of two FAb' fragments held together by two disulfide bonds; (5) Fv, a genetically engineered fragment containing the variable region of the light chain and the variable region of the heavy chain expressed as two chains; and (6) single chain antibody ("SCA"), a genetically engineered molecule containing the variable region of the light chain, the variable region of the heavy chain, linked by a suitable polypeptide linker as a genetically fused single chain molecule. Methods of making these fragments are routine.

Subject: Living multi-cellular vertebrate organisms, a category that includes both human and non-human mammals.

Transformed: A transformed cell is a cell into which has been introduced a nucleic acid molecule by molecular biology techniques. As used herein, the term transformation encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with viral vectors, transformation with plasmid vectors, and introduction of naked DNA by electroporation, lipofection, and particle gun or other acceleration techniques (for example, air gun).

Vaccine: A term used herein to mean a composition useful for stimulating a specific immune response (or immunogenic response) in a vertebrate. In some embodiments, the immunogenic response is protective or provides protective immunity, in that it enables the vertebrate animal to better resist infection with or disease progression from the organism against which the vaccine is directed. Without wishing to be bound by theory, it is believed that an immunogenic

response may arise from the generation of neutralizing antibodies, T-helper, or cytotoxic cells of the immune system, or all of the above.

In some embodiments, an "effective amount" or "immune-stimulatory amount" of a vaccine or vaccinating composition is an amount which, when administered to a subject, is sufficient to engender a detectable immune response. Such a response may comprise, for instance, generation of an antibody specific to one or more of the epitopes provided in the vaccine. Alternatively, the response may comprise a T-helper or CTL-based response to one or more of the epitopes provided in the vaccine. All three of these responses may originate from naïve or memory cells. In other embodiments, a "protective effective amount" of a vaccine or vaccinating composition is an amount which, when administered to a subject, is sufficient to confer protective immunity upon the subject.

Vector: A nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid sequences that permit it to replicate in a host cell, such as an origin of replication. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

Virus: Microscopic infectious organism that reproduces inside living cells. A virus typically consists essentially of a core of a single nucleic acid surrounded by a protein coat, and has the ability to replicate only inside a living cell. "Viral replication" is the production of additional virus by the occurrence of at least one viral life cycle. A virus may subvert the host cells' normal functions, causing the cell to behave in a manner determined by the virus. For example, a viral infection may result in a cell producing a cytokine, or responding to a cytokine, when the uninfected cell does not normally do so.

"Retroviruses" are RNA viruses wherein the viral genome is RNA. When a host cell is infected with a retrovirus, the genomic RNA is reverse transcribed into a DNA intermediate which is integrated very efficiently into the chromosomal DNA of infected cells. The integrated DNA intermediate is referred to as a provirus. The term "lentivirus" is used in its conventional sense to describe a genus of viruses containing reverse transcriptase. The lentiviruses include the "immunodeficiency viruses" which include human immunodeficiency virus type 1 and type 2, simian immunodeficiency virus, and feline immunodeficiency virus.

HIV is a retrovirus that causes immunosuppression in humans (HIV disease), and leads to a disease complex known as the acquired immunodeficiency syndrome. "HIV disease" refers to a well-recognized constellation of signs and symptoms (including the development of opportunistic infections) in persons who are infected by an HIV virus, as determined by antibody detection using ELISA or western blot studies. Alternatively, HIV infection can be detected by the presence of HIV RNA (for example, using RT-PCR) or HIV integrated DNA (for example, using PCR). Laboratory findings associated with this disease are a progressive decline in T-helper cells and a rise in viremia (viral load as determined by, for instance, RT-PCR).

Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. Hence "comprising A or B" means including A, or B, or A and B. For example, the term "a cell" includes a plurality of cells, including mixtures of two or more types of cells. The term "a nucleic acid molecule" includes a plurality of nucleic acid molecules, or a mixture of different nucleic acid molecules. Similarly, the same holds for "a protein" or "a polypeptide."

As used herein, the term "comprising" shall mean that the compositions and methods include the recited elements, but do not exclude other elements. "Consisting essentially of" shall mean excluding other elements of any essential significance to the combination. Thus, a composition consisting essentially of the elements as defined herein would not exclude trace contaminants from the isolation and purification method and/or pharmaceutically acceptable carriers, such as phosphate buffered saline, preservatives, and the like. "Consisting of" shall mean excluding more than trace elements of other ingredients and/or excluding substantial additional method steps. Embodiments defined by each of these transition terms are within the scope.

It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides, are approximate and are provided for description. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including explanations of terms, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

III. Overview of Several Embodiments

Provided herein in various embodiments are multi-clade, multivalent recombinant polyepitope polypeptides, which are useful to induce immunogenic responses in vertebrate animals to HIV-1. These polypeptides include CTL-stimulatory epitopes, T-helper cell stimulatory epitopes, B-cell stimulatory epitopes, or combinations of two or more such types of epitopes. Epitopes in the polyepitope polypeptides are selected to provide multi-clade coverage. In particular, epitopes are selected to be at least 50% conserved across a plurality of HIV-1 subtypes, for instance, at least 2, 3, 4, 5, 6, or more HIV-1 subtypes. In particular embodiments, at least 30% of the epitopes included in a single polyepitope polypeptide are at least 60% conserved, at least 70% conserved, at least 80% conserved, or even more highly conserved across a plurality of HIV-1 subtypes.

In specific embodiments there are provided isolated polyepitope polypeptides, wherein adjacent polypeptide segments are linked by a spacer peptide. In some examples, the spacer peptide

links multiple groups of polypeptide segments. Specific, non-limiting examples of the spacer peptide include the tri-amino acid lysine - alanine - alanine, or proline - glycine - proline.

In other examples, the isolated polyepitope polypeptides also include a targeting signal that targets the polyepitope polypeptides to a lysosome or to a proteosome. Specific, non-limiting, examples of the targeting signal include a targeting-competent fragment of lysosomal integral membrane protein-II or ubiquitin.

In still other examples, the isolated polyepitope polypeptides also include a plurality of amino acid segments from one or more HIV-1 coreceptors. A specific, non-limiting, example of a HIV-1 coreceptor is CCR5.

In further examples, the isolated polyepitope polypeptides include human cytotoxic T-lymphocyte stimulatory epitopes, human T-helper cell stimulatory epitopes, human B-cell stimulatory epitopes, or combinations of two or more epitopes thereof.

In additional embodiments there are provided isolated polyepitope polypeptides, which polypeptides comprise an amino acid sequence selected from the group consisting of sequences provided in SEQ ID NOs: 2, 4, 5, 6, 8 and 10. Also provided are mixtures of two or more isolated polyepitope polypeptides, including (but not limited to) mixtures of the polypeptides having sequences as shown in SEQ ID NOs: 2 and 8, 2 and 10, 4 and 8, 4 and 10, 5 and 8, 5 and 10, 6 and 8, and 6 and 10.

Other embodiments are isolated polynucleotides (nucleic acid molecules) which encode one of the polyepitope polypeptides described herein. Specific examples of such nucleic acid molecules comprise a sequence selected from the group consisting of sequences recited in SEQ ID NOs: 1, 3, 7, 9 and complements thereof. Other specific examples of nucleic acid molecules are the portions of each of SEQ ID NOs: 1, 3, 7, and 9 which correspond to and encode the polyepitope polypeptides shown in SEQ ID NOs: 2, 4, 8, and 10, respectively.

Also provided herein are genetic constructs that comprise at least one nucleic acid molecule encoding a polyepitope polypeptide, and host cells transformed with such a genetic construct.

Yet another embodiment is a composition comprising at least one polyepitope polypeptide or at least one nucleic acid molecule encoding a polyepitope polypeptide, and at least one component selected from the group consisting of pharmaceutically acceptable carriers and adjuvants. This disclosure further provides methods for eliciting and/or enhancing an immune response in a subject, which methods involve administering to the subject such a composition. In one specific, non-limiting example, the subject is infected with HIV-1

IV. Multi-Clade, Multivalent HIV-1 Constructs

The current disclosure provides multi-clade multivalent HIV-1 constructs useful for inducing immune responses in HIV-1-infected populations with diverse HLA alleles and HIV subtypes.

HIV-1 MCMV constructs comprise synthetic nucleic acid sequences to be used as HIV-1 immune-stimulatory and/or vaccine constructs to protect against multiple HIV-1 subtypes. These synthetic nucleic acid molecules or mixtures of nucleic acid molecules are designed to elicit both T-cell and B-cell responses against highly conserved epitopes within multiple HIV-1 subtypes. In specific embodiments, the synthetic genes are contained in plasmid constructs, which can be used as a DNA vaccine, as well as a source of recombinant protein for subsequent protein boosts.

Provided herein in various embodiments are multi-clade, multivalent polyepitope polypeptides, which are useful to induce immunogenic responses in vertebrate animals to HIV-1. These polypeptides include CTL-stimulatory epitopes, T-helper cell stimulatory epitopes, B-cell stimulatory epitopes, or combinations of two or more such types of epitopes. Epitopes in the polyepitope polypeptides are selected to provide multi-clade coverage. In particular, epitopes are selected to be at least 50% conserved across a plurality of HIV-1 subtypes, for instance, at least 2, 3, 4, 5, 6, or more HIV-1 subtypes. In particular embodiments, at least 30% of the epitopes included in a single polyepitope polypeptide are at least 60% conserved, at least 70% conserved, at least 80% conserved, or even more highly conserved across a plurality of HIV-1 subtypes.

One aspect of embodiments provided herein is that peptide sequences, each of which contains one or more antibody-binding or class I or class II MHC-binding epitopes, can be linked in tandem to form polyepitope polypeptides. These polyepitope polypeptides are proteolytically processed in cells to release the individual epitopes, and are useful for stimulating an immune response in a vertebrate animal.

When a MCMV polyepitope polypeptide is introduced into a cell, it is proteolytically processed into at least some of its constituent epitopes. At least some of the epitopes generated from the polypeptide can bind to MHC class I or MHC class II molecules present in the cell, though some of the epitopes may be specific for MHC class I or class II molecules present only on other cells. Included epitopes also may be B-cell epitopes, which elicit antibody-mediated immune responses upon binding to antibody receptors on the surface of a B-cell.

In one aspect, the disclosure features a nucleic acid encoding a polyepitope polypeptide that can include, in any order, a first, second, and third segment, each of which is at least nine amino acids in length. As used herein, a "segment" is an amino acid sequence which (a) corresponds to the sequence of a portion (that is, a fragment, less than all) of a naturally occurring protein, and (b) contains one or more epitopes. By "epitope" is meant a peptide which binds to the binding groove of an MHC class I or class II molecule, or to the antigen-binding region of an antibody. In addition, the polyepitope polypeptide can encode a targeting signal, for instance a peptide sequence that targets the protein to which it is fused to the lysosome or to the proteosome, as described in more detail herein.

In some embodiments, a segment of the polyepitope polypeptide has the amino acid sequence of a portion of (1) a naturally occurring HIV-1 protein or (2) a naturally occurring coreceptor (collectively referred to as "naturally occurring proteins"), that is at least nine amino acids in length. A second segment has the amino acid sequence of a second portion of the same or a

different naturally occurring protein, is at least nine amino acids in length, and includes at least one epitope which is different from the epitope(s) of the first segment. Optionally, a third segment is included in the polyepitope polypeptide, and has the amino acid sequence of a third portion of the same or a different naturally occurring protein, is at least nine amino acids in length, and includes at least one epitope which is different from the epitope(s) of the first and second segments. Alternatively, the first, second and third portions may be portions of two or three different naturally occurring proteins (for example, two or three different HIV-1 proteins). The polyepitope polypeptide may optionally include additional segments, for example, it can include at least 4, 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 90 or even 100 or more segments, each being a portion of a naturally-occurring protein of a pathogenic agent and/or of a naturally occurring coreceptor which can be the same or different from the protein(s) from which the other segments are derived.

Each of these segments is at least nine amino acids in length, and each contains at least one epitope different from the epitope(s) of other segments in the polyepitope polypeptide. At least one (and in some embodiments, more) of the segments in the polyepitope polypeptide may contain class I or class II MHC-binding epitopes. Two, three, or more of the segments can be contiguous in the polyepitope polypeptide: that is, they are joined end-to-end, with no spacer between them. Alternatively, any two adjacent segments can be linked by a spacer amino acid or spacer peptide. In particular embodiments, the spacer comprises three amino acids. Specific non-limiting examples of spacers are the tri-amino acid KAA and the tri-amino acid PGP. Additionally, a spacer amino acid or spacer peptide can be used to link multiple groups of two, three, or more contiguous segments in the polyepitope polypeptide: that is, a spacer amino acid or spacer peptide is inserted between every two, three, or more segments.

A given segment of protein within the polyepitope polypeptide need not be any specified length, so long as it is sufficiently long to generate at least one epitope, for example, 2, 3, 4, 5, or more epitopes, and is at least 9 amino acids in length. For example, a given segment can have a length of at least 10 amino acids, for example, at least 11, 12, 13, 14, 15, 20, 25, 30, 40, or 50 amino acids. A given segment corresponds to a particular naturally occurring protein if any 9 (or more) consecutive amino acids of the segment are found in exactly the same order in a portion of the naturally occurring protein. In exemplary embodiments, the segments included in a polyepitope polypeptide are obtained from one or more HIV-1 proteins and/or coreceptors (for example, CCR5).

It is understood that the term "naturally occurring proteins" used above does not preclude modification of the sequence used in the polyepitope polypeptide, for instance by changing one or a few amino acids. In addition, it is understood that the nucleic acid molecule encoding the segment need not be identical to the "naturally occurring" sequence, as found in (for instance) the HIV-1 genome. In particular, it is contemplated that the codon usage in the nucleic acid molecule can be modified, for instance to convert the encoding sequence to a codon optimized sequence. The codon optimization can be tailored for the host cell in which the construct will eventually be expressed. Thus, some constructs are engineered to be codon biased for expression in a prokaryotic cell, others

to be expressed in a unicellular eukaryotic cell, and still others to be expressed in a cell of a multicellular eukaryote (for example, a vertebrate). Codon selection to take advantage of species biases is well known to those of ordinary skill.

The discovery of the HIV-1 coreceptors, together with a greater understanding of the Envelope-receptor mediated conformational changes resulting in the membrane fusion process, has identified several promising vaccine targets. These epitopes as well as others in the transmembrane envelope glycoprotein (gp41) have been identified as HIV-1 neutralizing epitopes. Likewise, epitopes in the CCR5 coreceptor have been identified as potential targets for interfering with receptor-env interactions. Any of these epitopes can be included in the polyepitope polypeptides described herein.

Construction of HIV-MCMV Immunogens

HIV-1-MCMV immunogen constructs comprised of a string of codon-optimized epitopes have been produced. The antigenic fragments/epitopes in examples of such constructs were selected using published studies including broad MHC allele recognition and were compiled from the Los Alamos sequence database. A representative pair of immunogen constructs (polyepitope polypeptides) contains multiple B-cell epitopes, CTL epitopes, and T-helper epitopes representing immunodominant regions for all subtypes of HIV-1 (see tables included in the examples, and Figures 1, 2, and 3). The epitopes chosen are >80% homologous across diverse HIV-1 subtypes. B-cell epitopes in the virus binding domain of the human HIV coreceptor CCR5 are also included.

Without intending to be limited to a single interpretation, it is believed that antibodies to CCR5 together with neutralizing antibodies directed against the HIV-1 envelope glycoprotein and strong T-cell immunity will interfere with the viral entry process and is expected to induce sterilizing immunity.

Example immunogen constructs are shown in SEQ ID NOs: 2, 4, 5, 6, 8, and 10. The constructs shown in SEQ ID NOs: 2, 4, 5, and 6 include CTL epitopes (and therefore can be referred to generally as MCMV-CTL constructs); those in SEQ ID NOs: 8 and 10 include B-cell and T-helper epitopes (and therefore can be referred to generally as MCMV-AB/Th constructs).

Unique restriction enzyme digestion sites have been included in the nucleic acid constructs encoding the provided polyepitope polypeptides. These facilitate addition/deletion of epitopes, as well as the shuttling of the polyepitope cassette between a number of DNA vectors, including DNA vaccine constructs (for example, pVax-1, Invitrogen, Carlsbad, CA), eukaryotic yeast expression vectors (for example, pYes, Invitrogen, Carlsbad, CA), and multi-cell type expression vectors (for example, pTriEX-4, Novagen, Madison, WI). This enables the production of both a DNA based immunogen and vaccine, and ready production recombinant polyepitope polypeptide, which can be used directly as an immunogen or as a boost. The synthetic genes (which encode one or more polyepitope polypeptides) also can be incorporated into attenuated viral vectors such as Modified Vaccinia or Adenovirus to serve as a boosting agent.

Delivery and immunogenicity by inclusion of targeting sequences

Recent studies suggest that peptide spacers between epitopes and/or targeting sequences may increase the immunogenicity of certain epitopes. Targeting sequences such as the LIMP-II targeting sequence (which directs proteins to lysosomes and enhances class-II recognition), or targeting-competent fragments thereof, are used in certain provided embodiments to help enhance T-helper response. Likewise, proteosome targeting sequences (for example, ubiquitin or targeting-competent fragments thereof) that help induce class I recognition are included in specific embodiments, to provide improved CTL production. The chosen epitopes were back translated and human codon optimized for increased expression from the DNA construct.

In any of the described nucleic acids encoding polyepitope polypeptides, a spacer amino acid or spacer peptide can be included between any two adjacent segments of the construct. Optionally, in some embodiments the spacer is included between each epitope; in other embodiments, a spacer is included between every two, every three, every four, every five epitopes, or even less often. In particular embodiments, the spacer comprises three amino acids. Specific non-limiting examples of spacers are the tri-amino acid KAA and the tri-amino acid PGP.

Recognition of Epitopes contained in the Constructs

Most vaccine constructs under development are subtype-specific. This has led to development of a number of country-specific subtype-specific HIV-1 vaccines, however, such vaccines will be difficult to implement due to emerging diversity and changing epidemic of HIV-1.

In contrast, the constructs provided herein comprise highly conserved immunogenic regions of HIV-1 that result in cross-protective immune responses across HIV-1 subtypes. The immune responses to the immunogenic epitopes can be tested, for instance, in recently-infected HIV-1 infected persons (Primary HIV-1 infection; PHI) or individuals that have a slow progression to disease.

V. Uses of MCMV Immunogens

In order to function effectively *in vivo* as a DNA-based immunogen, it is advantageous to include within the MCMV nucleic acid construct a control sequence that has the effect of enhancing or promoting the translation of the sequences encoding the antigens. Use of such promoters is well known to those of skill in the fields of molecular biology, cell biology, and viral immunology (See, "Molecular Cloning: A Laboratory Manual", 2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; and "Current Protocols in Molecular Biology", Ausubel *et al.*, John Wiley and Sons, New York 1987 (updated quarterly)).

In certain embodiments, the nucleic acid construct is intended for use as a vaccine in a mammalian host. Therefore it is advantageous to employ a promoter which operates effectively in mammalian cells. Particular embodiments relate to both prokaryotic and eukaryotic host cells. Many

promoter sequences are known that are useful in either prokaryotic or eukaryotic host cells. A promoter is operably disposed with respect to the sequence(s) whose translation is to be promoted, so that it is capable of promoting translation. In certain embodiments, the promoter is the cytomegalovirus early promoter. In addition, in some embodiments, the sequences to be expressed are followed by a terminator sequence.

Preparation of the nucleic acids is readily accomplished by methods well known to workers of skill in the field of molecular biology. Procedures involved are set forth, for example, in Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989, and "Current Protocols in Molecular Biology", Ausubel *et al.*, John Wiley and Sons, New York 1987 (updated quarterly). Incorporation of promoters, such as the cytomegalovirus promoter, and of the polyadenylation signal, is likewise well known to skilled practitioners in molecular biology and recombinant DNA engineering.

When a nucleic acid molecule harboring a MCMV epitope chain is prepared, it may be obtained in larger quantities by methods that amplify a nucleic acid fragment. Such methods are widely known to workers skilled in molecular biology and recombinant DNA engineering. Examples of these methods include incorporation of the nucleic acid fragment into a plasmid for replication by culturing in a cell (for example, a prokaryotic cell) and harvesting the plasmid after growth of the culture, as well as amplification of the nucleic acid fragment by nucleic acid amplification methods, such as the PCR. These methods are exemplary only, and not intended to limit the ways in which the nucleic acid construct may be obtained.

The MCMV nucleic acid constructs may be introduced into appropriate host cells in many ways well known to those of ordinary skill in the fields of molecular biology and viral immunology. By way of example, these include, but are not limited to, incorporation into a plasmid or similar nucleic acid vector which is taken up by the host cells, or encapsulation within vesicular lipid structures such as liposomes, especially liposomes comprising cationic lipids, or adsorption to particles that are incorporated into the host cell by endocytosis.

In general, a host cell is a prokaryotic or eukaryotic cell harboring a MCMV nucleic acid, or into which such a MCMV molecule has been introduced. The constructs described herein induce the intracellular biosynthesis of the encoded multivalent HIV-1 antigens. A suitable host cell is one which has the capability for the biosynthesis of the gene products as a consequence of the introduction of the nucleic acid. In particular embodiments, a suitable host cell is one which responds to a control sequence and to a terminator sequence, if any, which may be included within the construct. In order to respond in this fashion, such a host cell contains within it components which interact with a control sequence and with a terminator, and act to carry out the respective promoting and terminating functions. When the host cell is cultured *in vitro*, it may be a prokaryote, a single-celled eukaryote or a vertebrate cell. In particular embodiments, the host cell is a mammalian cell

VI. Stimulation of Immunological Responses to HIV-1

With the provision herein of polyepitope polypeptide antigens specific to HIV-1, methods are now enabled for the stimulation of immune responses to such antigens in subjects. In certain embodiments, such immune responses will be protective against HIV-1 infection in the subject. MCMV polyepitope polypeptides (singly or in combination) can be used, for instance, as immunogenic agents in the inhibition, treatment, or amelioration of HIV-1. Subjects selected for this type of treatment are those who are known to have, or are suspected of having or are at risk of suffering, a HIV-1 infection.

The provided immunostimulatory MCMV polyepitope polypeptides, or constructs or vectors encoding such polypeptides, are combined with a pharmaceutically acceptable carrier or vehicle for administration as an immunostimulatory composition or a vaccine to human or animal subjects. In some embodiments, more than one polyepitope polypeptide may be combined to form a single preparation.

The immunogenic formulations may be conveniently presented in unit dosage form and prepared using conventional pharmaceutical techniques. Such techniques include the step of bringing into association the active ingredient and the pharmaceutical carrier(s) or excipient(s). In general, the formulations are prepared by uniformly and intimately bringing into association the active ingredient with liquid carriers. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring only the addition of a sterile liquid carrier, for example, water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets commonly used by one of ordinary skill in the art.

In certain embodiments, unit dosage formulations are those containing a dose or unit, or an appropriate fraction thereof, of the administered ingredient. It should be understood that in addition to the ingredients particularly mentioned above, formulations encompassed herein may include other agents commonly used by one of ordinary skill in the art.

The compositions provided herein, including those for use as immunostimulatory agents or vaccines, may be administered through different routes, such as oral, including buccal and sublingual, rectal, parenteral, aerosol, nasal, intramuscular, subcutaneous, intradermal, and topical. They may be administered in different forms, including but not limited to solutions, emulsions and suspensions, microspheres, particles, microparticles, nanoparticles, and liposomes.

The volume of administration will vary depending on the route of administration. By way of example, intramuscular injections may range from about 0.1 ml to about 1.0 ml. Those of ordinary skill in the art will know appropriate volumes for different routes of administration.

The amount of protein in each vaccine dose is selected as an amount that induces an immunostimulatory or immunoprotective response without significant, adverse side effects. Such amount will vary depending upon which specific immunogen is employed and how it is presented. Initial injections may range from about 1 μg to about 1 mg, with some embodiments having a range of about 10 μg to about 800 μg, and still other embodiments a range of from about 25 μg to about 500 μg. Following an initial vaccination, subjects may receive one or several booster immunizations, adequately spaced. Booster injections may range from about 1 μg to about 1 mg, with other embodiments having a range of about 10 μg to about 750 μg, and still others a range of about 50 μg to about 500 μg. Periodic boosters at intervals of 1-5 years, for instance three years, may be desirable to maintain the desired levels of protective immunity.

As described in WO 95/01441, the course of the immunization may be followed by *in vitro* proliferation assays of PBL (peripheral blood lymphocytes) co-cultured with ESAT6 or ST-CF, and especially by measuring the levels of IFN-released from the primed lymphocytes. The assays are well known and are widely described in the literature, including in U.S. Patent Nos. 3,791,932; 4,174,384 and 3,949,064.

A relatively recent development in the field of immune stimulatory compounds (for example, vaccines) is the direct injection of nucleic acid molecules encoding peptide antigens (broadly described in Janeway & Travers, *Immunobiology: The Immune System In Health and Disease,* page 13.25, Garland Publishing, Inc., New York, 1997; and McDonnell & Askari, *N. Engl. J. Med.* 334:42-45, 1996). Plasmids (vectors) that include nucleic acid molecules described herein, or that include a nucleic acid sequence encoding an immunogenic MCMV polyepitope polypeptide may be utilized in such DNA vaccination methods.

Thus, the terms "immunostimulatory preparation" and "vaccine" as used herein also include nucleic acid vaccines in which a nucleic acid molecule encoding a MCMV polyepitope polypeptide is administered to a subject in a pharmaceutical composition. For genetic immunization, suitable delivery methods known to those skilled in the art include direct injection of plasmid DNA into muscles (Wolff *et al., Hum. Mol. Genet.* 1:363, 1992), delivery of DNA complexed with specific protein carriers (Wu *et al., J. Biol. Chem.* 264:16985, 1989), co-precipitation of DNA with calcium phosphate (Benvenisty and Reshef, *Proc. Natl. Acad. Sci.* 83:9551, 1986), encapsulation of DNA in liposomes (Kaneda *et al., Science* 243:375, 1989), particle bombardment (Tang *et al., Nature* 356:152, 1992; Eisenbraun *et al., DNA Cell Biol.* 12:791, 1993), and *in vivo* infection using cloned retroviral vectors (Seeger *et al., Proc. Natl. Acad. Sci.* 81:5849, 1984).

Similarly, nucleic acid vaccine preparations can be administered via viral carrier.

It is also contemplated that the provided immunostimulatory molecules and preparations can be administered to a subject indirectly, by first stimulating a cell *in vitro*, which stimulated cell is thereafter administered to the subject to elicit an immune response.

VII. Immunological and Pharmaceutical Compositions

Immunological compositions, including immunological elicitor compositions and vaccines, and other pharmaceutical compositions containing latency-specific polypeptides or antigenic fragments thereof are useful for reducing, ameliorating, treating, or possibly preventing HIV infection, particularly HIV-1 infection. One or more of the polypeptides are formulated and packaged, alone or in combination with adjuvants or other antigens, using methods and materials known to those skilled in the vaccine art. An immunological response of a subject to such an immunological composition may be used therapeutically or prophylactically, and in certain embodiments provides antibody immunity and/or cellular immunity such as that produced by T-lymphocytes, such as cytotoxic T-lymphocytes or $CD4^+$ T-lymphocytes.

The MCMV polyepitope polypeptides may be administered with an adjuvant in an amount effective to enhance the immunogenic response against the conjugate. At this time, the only adjuvant widely used in humans has been alum (aluminum phosphate or aluminum hydroxide). Saponin and its purified component Quil A, Freund's complete adjuvant and other adjuvants used in research and veterinary applications have toxicities which limit their potential use in human vaccines. However, chemically defined preparations such as muramyl dipeptide, monophosphoryl lipid A, phospholipid conjugates such as those described by Goodman-Snitkoff *et al.* (*J. Immunol.* 147:410-415, 1991), encapsulation of the conjugate within a proteoliposome as described by Miller *et al.* (*J. Exp. Med.* 176:1739-1744, 1992), and encapsulation of the protein in lipid vesicles may also be useful.

The compositions provided herein, including those formulated to serve as vaccines, may be stored at temperatures of from about –100° C to about 4° C. They may also be stored in a lyophilized state at different temperatures, including higher temperatures such as room temperature. The preparation may be sterilized through conventional means known to one of ordinary skill in the art. Such means include, but are not limited to, filtration, radiation and heat. The preparations also may be combined with bacteriostatic agents, such as thimerosal (ethyl(2-mercaptobenzoate-S)mercury sodium salt) (Sigma Chemical Co., St. Louis, MO), to inhibit bacterial growth.

A variety of adjuvants known to one of ordinary skill in the art may be administered in conjunction with the protein(s) in the provided vaccine composition. Such adjuvants include but are not limited to the following: polymers, co-polymers such as polyoxyethylene-polyoxypropylene copolymers, including block co-polymers; polymer P1005; Freund's complete adjuvant (for animals); Freund's incomplete adjuvant; sorbitan monooleate; squalene; CRL-8300 adjuvant; alum; QS 21, muramyl dipeptide; CpG oligonucleotide motifs and combinations of CpG oligonucleotide motifs; trehalose; bacterial extracts, including mycobacterial extracts; detoxified endotoxins; membrane lipids; or combinations thereof.

In a particular embodiment, a vaccine is packaged in a single dosage for immunization by parenteral (that is, intramuscular, intradermal or subcutaneous) administration or nasopharyngeal (that is, intranasal) administration. In certain embodiments, the vaccine is injected intramuscularly into the deltoid muscle. The vaccine may be combined with a pharmaceutically acceptable carrier to

facilitate administration. The carrier is, for instance, water, or a buffered saline, with or without a preservative. The vaccine may be lyophilized for resuspension at the time of administration or in solution.

The carrier to which the polypeptide may be conjugated may also be a polymeric delayed release system. Synthetic polymers are particularly useful in the formulation of a vaccine to affect the controlled release of antigens.

Microencapsulation of the polypeptide will also give a controlled release. A number of factors contribute to the selection of a particular polymer for microencapsulation. The reproducibility of polymer synthesis and the microencapsulation process, the cost of the microencapsulation materials and process, the toxicological profile, the requirements for variable release kinetics and the physicochemical compatibility of the polymer and the antigens are all factors that must be considered. Examples of useful polymers are polycarbonates, polyesters, polyurethanes, polyorthoesters, polyamides, poly-(d,l-lactide-co-glycolide) (PLGA), and other biodegradable polymers.

Doses for human administration of a pharmaceutical composition or a vaccine may be from about 0.01 mg/kg to about 10 mg/kg, for instance about 1 mg/kg. Based on this range, equivalent dosages for heavier (or lighter) body weights can be determined. The dose may be adjusted to suit the individual to whom the composition is administered, and may vary with age, weight, and metabolism of the individual, as well as the health of the subject. Such determinations are left to the attending physician or another familiar with the subject and/or the specific situation. The vaccine may additionally contain stabilizers or physiologically acceptable preservatives, such as thimerosal.

The following examples are provided to illustrate certain particular features and/or embodiments. These examples should not be construed to limit the invention to the particular features or embodiments described.

EXAMPLES

Existing HIV-1 vaccine constructs are subtype specific. Though multiple sub-type specific candidate vaccines are under development, areas with high numbers of recombinant viruses would likely pose problems for subtype specific vaccines.

While the high degree of HIV variability has made vaccine design difficult, the proximity of populations with varying subtypes and the ease of travel have made a vaccine that can provide protection from multiple subtypes desirable. To address the problem of subtype variability, these examples illustrate production of constructs containing conserved B, T-helper, and CTL epitopes, with and without targeting domains. The chosen epitopes are expected to generate immune responses to multiple HIV-1 subtypes.

Selection of epitopes

In order to assemble a set of CTL epitopes that were conserved across a wide range of HIV-1 subtypes and that would be recognized by a large percentage of the population, the following databases and sources of sequence were consulted: Los Alamos HIV Molecular Immunology Database, Described Epitopes, LTNPs, EU and the Los Alamos HIV Sequence Database. The literature was also consulted, to locate reported conserved epitopes.

The epitopes were selected based on conserved epitopes previously shown to be recognized by HIV-1-infected persons from published reports or the Los Alamos Data base. These antigenic fragments/epitopes were selected using the results of *in vitro* and *in vivo* protection studies compiled in the Los Alamos database, as well as using the MotifScan software program. The following table (Table 1) provides the list of the multiple epitopes in the highly conserved regions in gag (p17, p24) pol (Prt, RT and Int) as well as Nef, Vif, Vpr, and Env epitopes selected based on their MHC class I binding (CTLs).

Epitopes were selected that were greater than 50% conserved across all the available sequences. Of these, the majority of epitopes were >90% conserved for subtypes A/B/C/D/E/F/G. Also taken into consideration was the frequency of Class I alleles that would recognize each conserved epitope.

Table 1:

CTL	Region	Sequence	AA Position in SEQ ID NO:2	HLA-alleles	Source
p17	18-29	KIRLRPGGKKKY	96-107	A3, A3.1, B27, B42, Bw62	Ferrari *et al.* ARHR 16:1433-1443, 2000
p17	36-44	WASRELERF	108-116	B35	
p17	74-92	ELRSLYNTVATLYCVH QRI	77-95	B8, A2 ,A11	
p24	15-27	ISPRTLNAWVKVV	120-132	A2, B57	
p24	30-55	KAFSPEVIPMFSALSEG ATPQDLNTM	133-158	B58, B44, B7	
p24	108-117	TSTLQEQIGW	159-168	B57	
p24	121-150	NPPIPVGDIYKRWIILG LNKIVRMYSPTSI	172-201	B8, B35, B27 ,B62, B52	Ferrari *et al.* ARHR 16:1433-1443, 2000, and Hanke *et al.*, Nature Med 9:951-955, 2000
p24	161-174	FRDYVDRFYKTLRA	202-215	B18, B44, A24, B70, B14	
p24	191-205	VQNANPDCKTILKAL	216-230	B51, B8,	Ferrari *et al.* ARHR 16:1433-1443, 2000
	216-226	ACQGVGGPGHK	231-241	A11	

pol	59-65	ITLWQRPLV	245-253	A28, A6802	Ferrari *et al.* ARHR 16:1433-1443, 2000; Hanke *et al.*, Nature Med 9:951-955, 2000
pol	262-273	TVLDVGDAYFSV	254-265	A2, A0201, B35	Ferrari *et al.* ARHR 16:1433-1443, 2000
pol	308-321	WKGSPAIFQSSMTK (SEQ ID NO:11)	N/A	B7, B35, A11, A3, A33	Ferrari *et al.* ARHR 16:1433-1443, 2000; Hanke *et al.*, Nature Med 9:951-955, 2000
pol	464-472	ILKEPVHGV	308-316	A2, A0201	Ferrari *et al.* ARHR 16:1433-1443, 2000
pol	495-507	QIYQEPFKNLKTG	320-332	A11	
pol	587-602	EPIVGAETFYVDGAAN	333-348	B35, B51, A28	
pol	956-964	LLWKGEGAV	358-366	A2, A0201	
pol		VIYQYMDDL	349-357	A0201	Hanke *et al.*, Nature Med 9:951-955, 2000
vif	48-57	HPKVSSEVHI	380-389	B0702	Altfeld *et al*, J. Immunology 167:2743-2752 (non-B sequence was derived from sequence data contained in the Los Alamos HIV sequence database)
Vif	17-26	RIRTWKSLVK	370-379	A0301	Altfeld *et al*, J. Immunology 167:2743-2752
vprB	29-42	AVRHFPRIWLHSL (SEQ ID NO:12)	N/A	B5701	
vpr nonB	29-42	AVRHFPRPWLHGL (SEQ ID NO:13)	N/A	B7301*	Altfeld *et al*, J. Immunology 167:2743-2752 (non-B sequence was derived from sequence data contained in the Los Alamos HIV sequence database)
vpr	58-66	AIIRILQQL	408-416	A0201	Altfeld *et al*, J. Immunology 167:2743-2752

nef	64-95	VGFPVRPQVPLRPMTY KGAVDLSHFLKEKGGL	420-451	A11, B8, B35, B7, A3, A2,	Ferrari *et al.* ARHR 2000 16:1433-1443, Hanke *et al.*, Nature Med 9:951-955, 2000
nef	127-141	GPGVRYPLTFGWCY	452-465	B57	
gp120	36-51	TVYYGVPVWKEATTT L	478-493	A3, B35, B55	Ferrari *et al.* ARHR 16:1433-1443, 2000
gp120	120-128	KTLPLCVTL	469-477	A2	
gp41	47-55	RAIEAQQHL	494-502	B51	
gp41		ERYLKDQQL (SEQ ID NO:14)	N/A	B14, A24,B8	Hanke *et al.*, Nature Med 9:951-955, 2000
SIV p27		ACTPYDINQML	291-301	MONKEY Mamu-A*01	
gp120		RGPGRAFVTI	278-287	Mouse H-2D	

These epitopes can be further characterized in summary as follows:

Table 2:

Epitope	# of Epitopes	HLA
p17	6	A3, A3.1, B27, B42, Bw62, B35, B8, A2, A11
p24	20	B57,A2,B58,B44,B7,B57,B8,B35,B27,B2,B52,B18,B44,A24,B 70,B14,B51,B8,B8,A11
pol	10	A28, B35, A2, A0201, B7, B35, A11, A3, A 2, A0201, A11, B35, B51, A28, A2, A0201
nef	8	A1, B8, B35, B7, B35, A3, A11, A2, B35, A11, B8, B57
gp120	4	A3, B55, B35, A2
gp-41	2	B51, B14
Vif	2	B0702, A0301
Vpr	3	B5701, B7301, A0201
SIVp27/gp120[1]	2	MAMU-A*01 / Mouse H-2D

[1] Control epitopes included for animal studies; SIVp27 is a simian epitope; the gp120 epitope is known to be recognized by the murine H-2D HLA.

CTL/Proteosome Constructs

Figure 1A shows the schematic map of one synthetic construct prepared to encode the identified conserved CTL epitopes. Representative nucleotide and amino acid sequences of HIV-1-MCMV-CTL with and without ubiquitin are shown in the accompanying Sequence Listing.

Using the 55 conserved CTL epitopes identified as described above, a synthetic gene was constructed using SOE followed by PCR. The synthesis is illustrated in Figure 2A. A tri-amino acid spacer (KAA) was inserted between every 3-5 epitopes, to enhance peptide processing. Two parallel constructs were constructed, one with and one without ubiquitin (included to provide proteosomal targeting and further enhance peptide processing). pVAXMCMV-CTL (Figure 2C) contains a 1.3

KB fragment coding for 55 CTL epitopes. pVAXMCMV-CTL-ubiquitin (Figure 2B) contains a 1.5 KB fragment coding for 55 CTL epitopes, covalently linked (5') to a modified ubiquitin molecule.

Expression of Multivalent Polypeptide from a MCMV Construct

Recombinant proteins encoded by the synthetic genes have been produced in *E. coli* (Figure 4). The expressed protein is larger, because the fusion tag attached to the protein is large. The fusion is detected (for example, in Figure 4) with an anti-Histidine monoclonal antibody, which detects the fusion tag (poly-His) encoded by the vector.

Protein expression of Ubiquitin-CTL gene also has been confirmed in HeLa cells. A typical western blot analysis of HeLa Cells transfected with 1 μg of pVAX MCMV-CTL-ubiquitin is shown in Figure 5. An anti-ubiquitin antibody was used for detection. These results indicate expression of an ubiquitinated protein of the correct predicted molecular weight as the synthetic CTL-ubiquitin gene. A band representing the normal cellular ubiquitin protein is also present on the blot.

Elispot Assays

To determine the biological relevance of epitopes used in the MCMV-CTL construct, epitopes contained in the construct were tested for recognition with PBMCs from HIV-1 infected individuals by doing Elispot assays. This example demonstrates that people infected with genetically different viruses could recognize epitopes included in the construct. Also since the individuals being tested are from different geographic locations, they are expected to have differences in their MHC molecules.

Epitope Testing (Elispot)

Chronically HIV-1-infected individuals infected with subtype B were selected from the HIV outpatient clinic at Johns Hopkins Hospital (Baltimore, MD) for testing cellular immune responses to HIV-1 (Keating *et al., AIDS Research and Human Retroviruses* 18:1067-1079, 2002). Subject median age was 42 years, with a range of 25-58 years. Median viral load was 2,228 copies/ml with a range of less than 330 copies/ml to 37,716 copies/ml. Two of the subjects (Nos. 10 and 15) used in the analysis had viral loads of greater than 15,000 copies/ml, and were identified as non-responders to subtype B Gag peptides in the test of Keating *et al.* (*AIDS Research and Human Retroviruses* 18:1067-1079, 2002). The average number of years patients were HIV-1 infected was 9.46 years (range 5-17 years), and median $CD4^+$ cell count was 534 cells/mm^3 with a range of 294-1009 cells/mm^3. These data represent a unique cohort of patients with strong immunological control as they have high CD4 values and have been HIV-1 infected for an average of nine years.

Peripheral blood was obtained by venipuncture and collected in heparin (Sigma, St. Louis, MO). PBMCs were isolated by Ficoll-Hypaque (Pharmacia-Amersham, Piscataway, NJ) gradient centrifugation, frozen in fetal calf serum (FCS; Summit) with 10% DMSO (Sigma, St. Louis, MO) and stored at -140°C.

A 9x9 matrix (shown below) representing the 55 epitopes contained in the construct plus 23 control peptides from Flu, EBV and CMV was generated to study Elispot responses, using 1 x 10^5 cryopreserved PMBCs per well and 5 μg/ml peptide (below). Samples were run in duplicate, and the positive cutoff defined as wells that had 2x more spots than negative control wells and at least 10 spots. Responses identified in the matrix testing were confirmed with individual peptide testing. Elispot was carried out essentially as described in Keating *et al., AIDS Research and Human Retroviruses* 18:1067-1079, 2002.

Table 3: Peptide matrix

	M1	M2	M3	M4	M5	M6	M7	M8	M9
MA	1-p17	2-p17	3-p17	4-p17	5-p17	6-p17	7-p24	8-p24	9-p24
MB	10-p24	11-p24	12-p24	13-p24	14-p24	15-p24	16-p24	17-p24	18-p24
MC	19-p24	20-p24	21-p24	22-p24	23-p24	24-p24	25-p24	26-p24	27-pol
MD	28-pol	29-pol	30-pol	31-pol	32-pol	33-pol	34-pol	35-pol	36-pol
ME	37-nef	38-nef	39-nef	40-nef	41-nef	42-nef	43-nef	44-nef	45-gp120
MF	46-gp120	47-gp120	48-gp120	49-gp41	50-gp41	51-vif	52-vif	53-vpr	54-vpr
MG	55-vpr	56-flu	57-flu	58-EBV	59-flu	60-CMV	61-flu	62-EBV	63-EBV
MH	64-EBV	65-EBV	66-EBV	67-Flu	68-EBV	69-CMV	70-EBV	71-EBV	72-EBV
MI	73-Flu	74-Flu	75-EBV	76-EBV	77-EBV	78-CMV			

Table 4: Peptides in the Matrix

No.	Peptide list	Sequences	AA Position in SEQ ID NO: 2	HLA
1	p17-1	KIRLRPGGK	96-104	A3, A3.1, B27
2	p17-2	RLRPGGKKKY	98-107	B42, Bw62
3	p17-3	WASRELERF	108-116	B35
4	p17-4	ELRSLYNTV	77-85	B8
5	p17-5	SLYNTVATL	80-88	A2
6	p17-6	TLYCVHQRI	87-95	A11
7	p24-1	ISPRTLNAW	120-128	B57
8	p24-2	TLNAWVKVV	124-132	A2
9	p24-3	KAFSPEVIPMF	133-143	B58
10	p24-4	IPMFSALSEGATPDQL (SEQ ID NO:15)	N/A	B44
11	p24-5	ATPQDLNTM	150-158	B7
12	p24-6	TSTLQEQIGW	159-168	B57
13	p24-7	NPPIPVGEIYKRWII (SEQ ID NO:16)	N/A	B8
14	p24-8	PPIPVGDIY	173-181	B35
15	p24-9	KRWIILGLNKIV	182-193	B27
16	p24-10	LGLNKIVRMYS	187-197	B62
17	p24-11	RMYSPTSI	194-201	B52
18	p24-12	FRDYVDRFYK	202-211	B18
19	p24-13	RDYVDRFYKTL	203-213	B44

20	p24-14	DYVDRFYKTL	204-213	A24
21	p24-15	YVDRFYKTL	205-213	B70
22	p24-16	DRFYKTLRA	207-215	B14
23	p24-17	VQNANPDCKTILKAL	216-230	B51
24	p24-18	NANPDCKTI	218-226	B8
25	p24-19	DCKTILKAL	222-230	B8
26	p24-20	ACQGVGGPGHK	231-241	A11
27	POL-1	ITLWQRPLV	245-253	A28
28	POL-2	TVLDVGDAY	254-262	B35
29	POL-3	VLDVGDAYFSV	255-265	A2, A0201
30	POL-4	WKGSPAIFQSSMT (SEQ ID NO:17)	N/A	B7, B35
31	POL-5	AIFQSSMTK (SEQ ID NO:18)	N/A	A11, A3
32	POL-6	ILKEPVHGV	308-316	A2, A0201
33	POL-7	QIYQEPFKNLKTG	320-332	A11
34	POL-8	EPIVGAETF	333-341	B35, B51
35	POL-9	AETFYVDGAAN	338-348	A28
36	POL-10	LLWKGEGAV	358-366	A2, A0201
37	NEF-1	VGFPVTPQVPLRPMT (SEQ ID NO:19)	N/A	A1, B8
38	NEF-2	FPVRPQVPL	422-430	B35
39	NEF-3	FPVRPQVPLR	422-431	B7
40	NEF-4	RPQVPLRPMTY	425-435	B35
41	NEF-5	QVPLRPMTYK	427-436	A3, A11, A2, B35
42	NEF-6	AVDLSHFLK	438-446	A11
43	NEF-7	FLKEKGGL	444-451	B8
44	NEF-8	GPGVRYPLTFGWCY	452-465	B57
45	gp120-1	TVYYGVPVWK	478-487	A3
46	gp120-2	VPVWKEATTT	483-492	B55,
47	gp120-3	VPVWKEATTTL	483-493	B35
48	gp120-4	KTLPLCVTL	469-477	A2
49	gp-41-1	RAIEAQQHL	494-502	B51
50	gp-41-1	ERYLKDGGL	503-511	B14
51	VIF-1	HPKVSSEVHI	380-389	B0702
52	VIF-2	RIRTTWKSLVK (SEQ ID NO: 20)	N/A	A0301
53	VPRB-1	AVRHFPRIWLHSL (SEQ ID NO: 21)	N/A	B5701
54	VPRNB-2	AVRHFPRPWLHGL (SEQ ID NO: 22)	N/A	B7301
55	VPR-3	AIIRILQQL	408-416	A0201
56	Influenza A PB1 591-599	VSDGGPNLY (SEQ ID NO: 23)	N/A	A1
57	Influenza A NP 44-52	CTELKLSDY (SEQ ID NO: 24)	N/A	A1
58	EBV BMLF 259-267	GLCTLVAML (SEQ ID NO: 25)	N/A	A2
59	Influenza A Matrix 1 58-66	GILGFVFTL (SEQ ID NO: 26)	N/A	A2
60	HCMV Pp65 495-503	NLVPMVATV (SEQ ID NO: 27)	N/A	A2
61	Influenza A NP 265-273	ILRGSVAHK (SEQ ID NO: 28)	N/A	A3

62	EBV BMLF 259-267	RVRAYTYSK (SEQ ID NO: 29)	N/A	A3
63	EBV EBNA3A 603-611	RLRAEAQVK (SEQ ID NO: 30)	N/A	A3
64	EBV EBNA3B 416-424	IVTDFSVIK (SEQ ID NO: 31)	N/A	A11
65	EBV BRLF1 134-143	ATIGTAMYK (SEQ ID NO: 32)	N/A	A11
66	EBV BRLF1 28-37	DYCNVLNKEF (SEQ ID NO: 33)	N/A	A24
67	Influenza A NP 91-99	KTGGPIYKR (SEQ ID NO: 34)	N/A	Aw68
68	EBV EBNA3A 379-387	RPPIFIRRL (SEQ ID NO: 35)	N/A	B7
69	HCMV Pp65 495-503	TPRVTGGGAM (SEQ ID NO: 36)	N/A	B7
70	EBV EBNA3A 158-166	QAKWRLQTL (SEQ ID NO: 37)	N/A	B8
71	EBV EBNA3A 325-333	FLRGRAYGL (SEQ ID NO: 38)	N/A	B8
72	EBV BZLF1 190-197	RAKFKQLL (SEQ ID NO: 39)	N/A	B8
73	Influenza A NP 380-388	ELRSRYWAI (SEQ ID NO 40)	N/A	B8
74	Influenza A NP 380-388	SRYWAIRTR (SEQ ID NO: 41)	N/A	B27
75	EBV EBNA3C 258-266	RRIYDLIEL (SEQ ID NO: 42)	N/A	B27
76	EBV EBNA3A 458-466	YPLHEQHGM (SEQ ID NO: 43)	N/A	B35
77	EBV EBNA3C 281-290	EENLLDFVRF (SEQ ID NO: 44)	N/A	B44
78	HCMV Pp65 495-503	QEFFWDANDIYRIFA (SEQ ID NO: 45)	N/A	B44

Results:

Eleven individuals chronically infected with HIV-1 subtype B (the subtype found in the US) were tested. Preliminary testing of individuals form Ivory Coast, West Africa (subtype A/G viruses) was also conducted. These data indicate that the CTL epitopes contained in the construct are recognized by individuals infected with genetically distinct subtypes of HIV-1.

The eleven patients had confirmed responses to multiple epitopes, in p17, p24, pol, vpr, gp41, gp120, and nef. Representative bar graphs demonstrating breadth and magnitude of CTL responses generated from PBMCs of six of the eleven individuals chronically infected with subtype/clade B HIV-1 are shown in Figure 6. These patient data are also summarized in the following table (which also includes data from the remaining five individuals), and indicate the viral load of the individual, their CD4 count, the known HLA type of the individual and the known HLA binding properties of the epitopes to which the individual responded.

Table 5:

Patient #	CD4 #	Viral load	Patient HLA	Response regions	Epitope HLA
3	501	11,694	A30, A33, B53, B14	p24-6, Nef-8	B57
7	349	6,177	NA	Pol-6	A2, A02201
10	700	37,716	NA	Pol-9	A28

11	924	1742	A33, A68, B7, B57	p24-1,p24-3, p24-4, VprB-1, Pol-9	B57, B58, B44, B35, B5701
14	843	<50	A2, A30, B13, B27	p24-9, VprNB-2	B27, B7301
15	349	20,354	NA	EBV Control only	B35
16	434	2714	A3, A26, B49, B65	p17-1, p24-16	A3, B27, B14
18	493	5862	A2, A29, B14, B72	p17-5, p24-5, p24-16	A2, B7, B14
19	822	1490	NA	p24-1, p24-3	B57, B58
21	567	<29	A3, A23, B35, B62	P17-1, p17-5 Nef-8 gp41-1, gp41-2	A3, B27, A2, B57, B14
22	1009	<50	A23, B57, B72	p24-1, p24-3, p24-6, Nef-8, gp120-1	B57, B58, B57, B57
Neg Control	NA	Uninfected	NA	EBV Control only	B8

Ten of the eleven (10/11) patients (90.9 %) responded to one or more peptides; 8/11 (72.7%) responded to two or more peptides in the multivalent construct. A summary of the peptides recognized from each gene region by the 11 chronically infected individuals (Subtype B) is provided in Table 6.

Table 6:

Gene Region	Peptides Recognized	Percentage
All peptides	18/55	32.7
Gag	10/26	38.4
Pol	2/10	20
Nef	1/8	12.5
Env	3/6	50
Vif	0/2	0
Vpr	2/3	66

Of the seven chronically HIV-infected study subjects who responded to HIV-1 individual peptides and had HLA typing made available, HLA specificities of the CTL epitopes were compared with the patient HLA types. All seven individuals responded to at least one of the predicted CTL epitopes according to the restricting HLA allele of that epitope. There was great variability as to whether patient cells could target those CTL epitopes predicted. Figure 7 shows the percentage of predicted epitopes that were targeted by the patients' $CD8^+$ cells in the Elispot assay. Patient number 21 responded to only one of the 12 peptides predicted to be targeted according to the patient's HLA type, whereas patient number 22 recognized all three of the peptides predicted to elicit Elispot responses. All of the 7 subject studies were capable of targeting epitopes outside of their respective HLA type. For example, patient number 3 who had been characterized for A30, A33, B53, and B14

HLA type, only recognized p24 peptide sequence TSTLQEQIGW, which is a B57 HLA restricted CTL epitope. Of the 7 peptides targeted by patient number 11, four were epitopes restricted to mismatched HLA haplotype. All five of the HLA restricted epitopes targeted by patient number 21 were disparate for that individual's HLA type.

Preliminary testing of individuals infected with other subtypes of virus (A from Kenya, C from India and A/E from Thailand) is underway.

Additional testing was carried out using PBMCs from HIV-1 positive blood donors from Ivory Coast. The testing was done as described for the subtype B individuals, in that the same peptide pool matrix was tested as describe above. The four Ivorian samples tested were from "healthy" blood donors and thus were presumed to be incident HIV-1 infections. Due to the high prevalence of recombinant subtype A/G viruses in this region, it is presumed these individuals were infected with A/G viruses. The four individuals had broad responses based on the results of the peptide matrix screen.

Individual 1 had predicted responses from the matrix to 46 peptides contained in MCMV. The gene regions represented by these peptides included p24, pol, nef, gp120, vpr, and vif. Individual 2 had predicted responses to 15 peptides (in pol, nef, gp41, and vif). Individual 3 had predicted responses to 10 peptides (in p24, nef, vpr, and gp120). Individual 4 had predicted responses to 47 epitopes contained in the following HIV-1 gene regions: p17, p24, pol, nef, gp120, gp41, vpr, and vif.

Overall, good CTL responses were observed to chosen epitopes (in gag, pol, env, nef, vpr, and vif) in subtype B infected individuals from the United States and presumed subtype A/G infected individuals from Ivory Coast.

Mouse studies

Transgenic HLA mice studies can be carried out to detect immunologic responses induced by each construct (with and without ubiquitin). A comparison of the with and without ubiquitin constructs will enable characterization of the effects of ubiquitin on epitope processing and immunogenicity.

One HIV-1 epitope was included in the MCMV-CTL construct (located in the middle of the construct) which has previously been shown to be recognized by mice expressing MHC class I H-2D. The inclusion of this epitope allows limited immunogenicity studies with any mouse strain that expresses the H-2D allele. In addition, transgenic mice that express human MHC molecules (such as the C57BL/6-TgN(HLA-A2.1) strain) can be used to further look at all of the epitopes in the construct that are A2.1 restricted. By way of example, 1-3 doses of the DNA can be tested, likely 2-5 μg of DNA at a time, applying the DNA interdermally using a gene gun.

Mice will be sacrificed and splenocytes will be harvested 7-10 days after the last injection. The splenocytes will then be used in Elispot assays to determine if the mice recognize specific peptides contained within the construct. Parallel studies can be done with the construct with and

without ubiquitin to demonstrate that the ubiquitin fusion increases CTL responses (breadth, A2.1 restricted epitopes, and magnitude).

Primate Studies

Due to the inclusion of the Mamu- A*01 restricted epitope, monkeys that have this HLA can be used to study the effects of adding the ubiquitin tag, and to characterize the resultant changes in the magnitude of the immunogenic responses.

B-CELL/T-HELPER CELL/LYSOSOME CONSTRUCTS

The B-cell (Table 7) and T-helper (Table 8) epitopes were chosen by literature searches and information and software contained in the Los Alamos HIV Molecular Immunology database.

Table 7:

B epitopes	Clade	Region	Sequence	AA Position in SEQ ID NO:8	Source
Tat	A	21-40	PCNKCYCKKCCYHCQVCFLN	79-98	Boykins *et al.* 2001 peptides 21: 1839/database
Tat	B	21-40	ACTNCYCKKCCFHCQVCFTT	2-21	
Tat	C	21-40	ACNTCYCKKCSYHCLVCFQT	146-165	
Tat	D	21-40	PCNKCYCKKCCYHCQVCFIT (SEQ ID NO: 46)	N/A	
Tat	A/E	21-40	ACSKCYCKKCCWHCQLCFLK (SEQ ID NO: 47)	N/A	
Tat	F	21-40	PCTKCYCKRCCFHCQWCFIT (SEQ ID NO: 48)	N/A	
Tat	A/G	21-40	ACSKCYCHICCWHCQLCFLN (SEQ ID NO: 49)	N/A	
Tat	A	53-68	KQRRGTPQSNKDHQNP	102-117	
Tat	B	53-68	RQRRRAPQDSQTHQVS	26-41	
Tat	C	53-68	RQRRSAPPSSEDHQNL	170-185	
Tat	D	53-68	RQRRRPPQGGQAHQDP (SEQ ID NO: 50)	N/A	
Tat	A/E	53-68	KHRRGTPQSSKDHQNP (SEQ ID NO: 51)	N/A	
Tat	A/G	53-68	RRRRGTPQSRQDHQNP (SEQ ID NO: 52)	N/A	
Tat	F	53-68	RQRHRTPQSSQIHQDP (SEQ ID NO: 53)	N/A	
gp120			HERSYMFSDLENRCI	214-228	2001 Vaccine meeting #295 Menendez *et al.*
gp41	A	2F5-4E10	NEQDLLALDKWANLWNWFDIS	122-142	Parker *et al.* J Virol. 2001 75:10906 Non-B subtype peptide sequences were determined using the Los Alamos HIV sequence Database subtype consensus sequence data.
gp41	B	2F5-4E10	NEQELLELDKWASLWNWFDIT	189-209	
gp41	C	2F5-4E10	NEKDLLALDKWQNLWSWFDIT	229-249	
gp41	D	2F5-4E10	NEKELLELDKWASLWNWFSIT (SEQ ID NO: 54)	N/A	
gp41	F	2F5-4E10	NEQELLALDKWASLWNWFDIS (SEQ ID NO: 55)	N/A	
gp41	G	2F5-4E10	NEQDLLALDKWASLWTWFSIT	N/A	

			(SEQ ID NO: 56)		
gp41		N1	SGIVQQQNNLLRAIEAQQHLLQ LTVWGIKQLQARIL (SEQ ID NO: 57)	N/A	Rosny *et al.* J Virol 2001 75:8859-8863
gp41		C1	WMEWDREINNYTSLIHSLIEES QNQQEKNEQELL	297-330	
human CCR5		ECL1 89-102	YAAAQWDFGNTMCQL (SEQ ID NO: 58)	N/A	Barassi *et al.* AIDS Vaccine 2001 abstract # 112 / Philadelphia, PA September 5-8, 2001
human CCR5		ECL2 178-197	CSSHFPYSQYQFWKNFQTLK (SEQ ID NO: 59)	N/A	

Table 8:

T-helper	Region	Sequence	AA Position in SEQ ID NO: 8	Source
p24	111-132	LQEQIGWMTNNPPIPVGEIYKR	386-407	Wilson *et al.* J Virol 2001 75:4195 and Cosimi and Rosenberg, Los Alamos HIV Molecular Immunology Database, 2000
p24	131-152	KRWIILGLNKIVRMYSPTSILD	406-427	Wilson *et al.*, J. Virol 2001 75:4195
p24	146-160	SPVSILDIRQGPKEP (SEQ ID NO: 60)	N/A	Cosimi and Rosenberg , Los Alamos HIV Molecular Immunology Database, 2000
p24	1-22	PIVQNIQGQMVHQAISPRTLNA	360-381	
p24	156-170	GPKEPFRDYVDRFYK	431-445	
p24	31-52	AFSPEVIPMFSALSEGATPQDL	338-359	
pol	36-52	EICTEMEKEGKISKIGP	446-462	
pol (rt)	303-317	FRKYTAFTIPSINNE	467-481	Wilson *et al.*, J. Virol 2001 75:4195
pol (rt)	335-349	SPAIFQSSMTKILEP	482-496	
pol (rt)	596-610	WEFVNTPPLVKLWYQ	497-511	
pol (int)	915-929	KTAVQMAVFIHNFKR	512-526	
pol (int)	956-970	QKQITKIQNFRVYYR	527-541	
vpr	66-80	QLLFIHFRIGCRHSR (SEQ ID NO: 61)	N/A	Cosimi and Rosenberg, Los Alamos HIV Molecular Immunology Database, 2000
rev	9-23	DEELIRTVRLIKLLY (SEQ ID NO: 62)	N/A	
rev	41-56	RRRRWRERQRQIHSIS (SEQ ID NO: 63)	N/A	
env	476-490	DMRDNWRSELYKYKV	596-610	
env	562-576	QQHLLQLTVWGIKQL	611-625	
env	667-681	ASLWNWFDITNWLWY	626-640	
env	682-696	IKIFIMIVGGLIGLR	641-655	
env	827-841	HIPRRIRQGLERALL (SEQ ID NO: 64)	N/A	

Construction of a MCMV construct containing these B-cell and T-helper epitopes was carried out essentially similarly to the procedures used to generate the MCMV-CTL construct. Representative sequences of MCMV-AB/Th construct are shown in SEQ ID NOs: 7 and 9; the encoded multivalent antigen polypeptides are shown in SEQ ID NOs: 8 and 10. A tri-amino acid spacer (GPG) was inserted between each of the Ab epitopes and between every 3-5 T-helper epitopes,

to provide additional flexibility in the molecule, and to enhance peptide presentation. In addition, the LIMP-II lysosomal targeting sequence was included at the C-terminus in one construct (SEQ ID NOs: 7 and 8), to enhance processing of the epitopes.

Synthetic peptides of the T-helper epitopes have been synthesized, using standard peptide synthesis protocols, for use in lymphocyte proliferation assays.

Targeting to Lysosome

Detection of targeting of a MCMV-AB/Th polypeptide to the lysosome can be accomplished using confocal microscopy. By way of example, HeLa cells can be transfected with the pVax-1 MCMV constructs (with and without and ubiquitin) followed by staining protocols to detect the lysosome (detection of LAMP-1 using a LAMP-1-specific antibody) and the expressed MCMV protein (for example, using polyclonal antibody generated by injection of the recombinant MCMV protein into mice). A detailed protocol for the detection of lysosomal targeting is found in Rodriguez *et al., J. Virology* 75:10421-10430, 2001.

Confirmation of Immunogenicity of Ab and T-helper construct(s)

In order to confirm the immunogenicity of the epitopes in a MCMV-AB/Th construct, T-helper assays (Lymphocyte proliferation assays) can be performed using PBMCs from HIV-1 infected individuals, using methods basically as described in Wilson *et al. J. Virology* 75:4195-4207, 2001.

For the antibody epitopes, sera/plasma from infected individuals can be used to test for the presence of antibodies that would react with the protein encoded by this construct. Additionally, mouse or monkey immunization studies with either the DNA construct or purified recombinant MCMV-AB/Th protein (practically any strain of mouse or primate routinely available) can be preformed to ascertain the production of antibodies. Animals would be injected with 1-3 doses of DNA (2-5 μg DNA for mice and 1-2 mg DNA for rhesus macaques) or purified recombinant protein (20-50 μg for mice or 50-500 μg for monkeys). Prior to the first injection and 1-2 weeks following each injection, blood will be drawn and tested for the presence of antibody specific to the MCMV-AB/Th epitopes by ELISA.

Other methods for testing the immunogenicity of Ab and T-helper epitopes will be known to those of ordinary skill in the art.

Determination of the optimal time frame for vaccination with the recombinant protein as an immunologic "boost"

Recombinant protein produced from the MCMV constructs described herein can be utilized in conjunction with the DNA immunogen(s), or other currently available DNA vaccines, as an immunological "boost".

Following the initial animal injections with the DNA construct(s) immune responses will be monitored (for instance, using CTLs, Elispot assays, T-helper/lymphocyte proliferation assays, and/or ELISA assays) to determine the peak of the immune response for each arm of the immune system (T-Cell and B-cell). Based on the observed responses, a series of boost injections of a MCMV polypeptide can be initiated. By studying the responses, the time frame to generate the maximum response from memory T or B-cells can be optimized. Systems for optimizing the boost effect will be known to those of ordinary skill in the art.

Clinical trials

Following the production of the vaccine materials, Phase I safety trials can be performed in populations at risk for HIV. In the United States, target populations would include, for example, gay male cohorts or IV Drug using cohorts. In countries other than the United States, potential populations would include, for example, prenatal cohorts, IV drug use cohorts and prostitutes.

It will be apparent that the precise details of the constructs, compositions, and methods described herein may be varied or modified without departing from the spirit of the described invention. We claim all such modifications and variations that fall within the scope and spirit of the claims below.